Joint confidence regions for the regression parameters

Stat 524
10-17-17

①

Assuming that $\vec{\varepsilon}$ has a multivariate normal distribution

$$\hat{\beta} \sim N_p(\beta, \sigma^2(X'X)^{-1})$$

so $\hat{\beta} - \beta \sim N_p(\vec{0}, \sigma^2(X'X)^{-1})$

and $(\hat{\beta}-\beta)'(\sigma^2(X'X)^{-1})^{-1}(\hat{\beta}-\beta) \sim \chi^2_p$

②

$$(\hat{\beta}-\beta)'\frac{X'X}{\sigma^2}(\hat{\beta}-\beta) \sim \chi^2_p \quad \star$$

Also, $\dfrac{(n-p)MSE}{\sigma^2} \sim \chi^2_{n-p}$ $\quad \star\star$

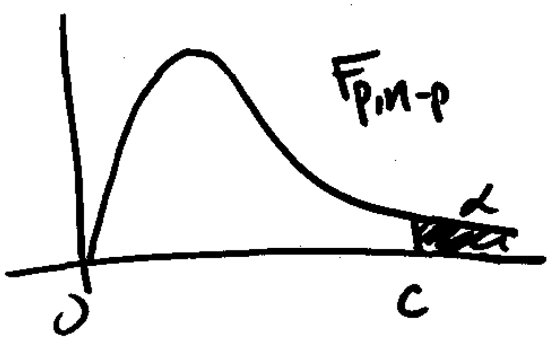
And $\star$ is independent of $\star\star$

So

$$\frac{(\hat{\beta}-\beta)'\frac{X'X}{\cancel{\sigma^2}}(\hat{\beta}-\beta)/p}{\frac{\cancel{(n-p)}MSE}{\cancel{\sigma^2}}\Big/\cancel{(n-p)}} \sim F_{p,n-p}$$

3

$$\therefore \quad \frac{(\hat{\beta}-\beta)' X'X (\hat{\beta}-\beta)}{p \, MSE} \sim F_{p,n-p}$$



$$P\left[\frac{(\hat{\beta}-\beta)' X'X (\hat{\beta}-\beta)}{p \, MSE} \le c \right] = 1-\alpha$$

4

Scheffé:

$(\hat{\beta}-\beta)' X'X (\hat{\beta}-\beta) \le c \, p \, MSE$ is an event whose probability is $1-\alpha$

Apply this to simple linear regression:

$$X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}_{n\times 2} \qquad X'X = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix}_{2\times n} \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}_{n\times 2}$$

$$X'X = \begin{bmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{bmatrix}_{2\times 2}$$

5

$$(\hat\beta-\beta)'X'X(\hat\beta-\beta) =$$

$$\begin{bmatrix}\hat\beta_0-\beta_0, & \hat\beta_1-\beta_1\end{bmatrix}\begin{bmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2\end{bmatrix}\begin{bmatrix}\hat\beta_0-\beta_0 \\ \hat\beta_1-\beta_1\end{bmatrix}$$

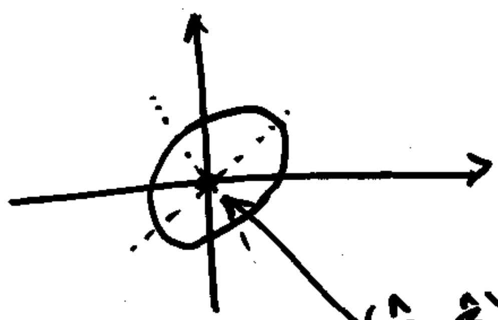
$$n(\hat\beta_0-\beta_0)^2 + 2\Sigma x(\hat\beta_0-\beta_0)(\hat\beta_1-\beta_1) + \Sigma x^2(\hat\beta_1-\beta_1)^2$$

set $\le c\,p\,MSE$

$$a x^2 + bxy + cy^2 \le k$$

This is the interior of an ellipse centered at $(\hat\beta_0, \hat\beta_1)$

6



$(\hat\beta_0, \hat\beta_1)$

The individual Scheffé confidence intervals are the projections of the ellipse (ellipsoid) onto the individual axes.

⑦

Result: Joint confidence intervals for $\beta_0, \beta_1, \ldots, \beta_k$:

$$\hat{\beta}_i \pm \Delta \, S.E.(\hat{\beta}_i)$$

Bonferroni: $\Delta = t_{n-p}\left(\frac{\alpha}{2p}\right)$

Scheffé : $\Delta = \sqrt{p F_{p,n-p}(\alpha)}$

⑧

Two special cases of regression

① What happens if the variables are all standardized ?

In simple linear regression, let

$$z_i = \frac{x_i - \bar{x}}{s_x} \quad \text{predictor}$$

and $$w_i = \frac{y_i - \bar{y}}{s_y} \quad \text{outcome}$$

(9)

Let b_1 and b_0 be the L.S. slope & intercept

$$b_1 = \frac{S_{zw}}{S_{zz}} \quad \text{and} \quad b_0 = \bar{w} - b_1 \bar{z} = 0 - 0 = 0$$

$$\begin{aligned} S_{zz} &= \sum z_i^2 - n\cancel{\bar{z}}^2 \\ &= \sum \left(\frac{x_i - \bar{x}}{s_x}\right)^2 = \frac{1}{s_x^2} \sum (x_i - \bar{x})^2 \\ &= \frac{S_{xx}}{S_{xx}/n-1} = n-1 \end{aligned}$$

(10)

$$\begin{aligned} S_{zw} &= \sum z_i w_i - n \cancel{\bar{z}\bar{w}} \\ &= \sum \left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right) \\ &= \frac{1}{s_x s_y} \underbrace{\sum (x_i - \bar{x})(y_i - \bar{y})}_{S_{xy}} \\ &= (n-1) \frac{s_{xy}}{s_x s_y} = (n-1) r \end{aligned}$$

so $b_1 = \dfrac{S_{zw}}{S_{zz}} = \dfrac{(n-1)r}{(n-1)} = r$

(11)

Summary: If x & y are standardized prior to fitting the regression line, then the slope will be r and the intercept will be 0.

$$\frac{y_i - \bar{y}}{s_y} = r\left(\frac{x_i - \bar{x}}{s_x}\right)$$

(12)

(2) Generalized least squares

Instead of assuming $\vec{\varepsilon} \sim N(\vec{0}, \sigma^2 I)$,

Assume $\vec{\varepsilon} \sim N(\vec{0}, \sigma^2 V)$

all elements of V are known and V is symmetric and <u>positive</u> <u>definite</u>

all eigenvalues are > 0

Since V is symmetric,

$$V = P \Lambda P', \text{ where } \Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

Let $$\Lambda^{-1/2} = \begin{bmatrix} \lambda_1^{-1/2} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1/2} \end{bmatrix}$$

Let $L = \Lambda^{-1/2} P'$

Our original model is $Y = X\beta + \varepsilon$

Now $LY = LX\beta + L\varepsilon$

14

$$\underbrace{LY}_{W} = \underbrace{LX}_{Z}\beta + \underbrace{L\varepsilon}_{\gamma}$$

$W = Z\beta + \gamma$ Does this fit our usual assumptions?

$$E[\gamma] = E[L\varepsilon] = L\,E[\varepsilon] = L\,\vec{0} = \vec{0} \checkmark$$

$$Cov[\gamma] = Cov[L\varepsilon] = L\,Cov(\varepsilon)L'$$

$$= \Lambda^{-1/2} P' \sigma^2 V (\Lambda^{-1/2} P')'$$

$$= \sigma^2 \Lambda^{-1/2} \cancel{P'(P} \Lambda \cancel{P')P} \Lambda^{-1/2}$$

$$= \sigma^2 I$$

Since the transformed model satisfies our usual assumptions,

$$\hat{\beta} = (Z'Z)^{-1} Z'W$$

$$= [(LX)'LX]^{-1} (LX)'LY$$

16

$$= (X'\underbrace{L'L}X)^{-1} X'\underbrace{L'L}Y$$

$$L'L = (\Lambda^{-1/2} P')' \Lambda^{-1/2} P'$$

$$= P \Lambda^{-1/2} \Lambda^{-1/2} P'$$

$$= P \Lambda^{-1} P'$$

$$= V^{-1}$$

$$\therefore \hat{\beta} = (X'V^{-1}X)^{-1} X'V^{-1}Y$$

This is the <u>generalized</u> <u>least</u> <u>squares</u> solution

From now on, our original solution of $\hat{\beta} = (X'X)^{-1}X'Y$ will be called the ordinary least squares solution